FORMULA SHEET
MATH 5651
SPRING 2007

INSTRUCTOR: PROF. GREG W. ANDERSON

Basic rules

- $0 \leq \Pr(A) \leq 1$.
- $\Pr(S) = 1$ (S: sample space).
- $\Pr(\emptyset) = 0$ ($\emptyset$: empty set).
- $\Pr(A_1 \cup \cdots \cup A_n) = \Pr(A_1) + \cdots + \Pr(A_n)$ for disjoint events $A_1, \ldots, A_n$.
- $\Pr(A^c) = 1 - \Pr(A)$ (A^c: complement of A).
- $A \subset B \Rightarrow \Pr(A) \leq \Pr(B)$.
- $\Pr(A) = \Pr(AB) + \Pr(AB^c)$.
- $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(AB)$.
- $\Pr(A \cup B \cup C) = \Pr(A) + \Pr(B) + \Pr(C)$
  $- \Pr(AB) - \Pr(BC) - \Pr(AC) + \Pr(ABC)$.

Combinatorial definitions and formulas

$$
\begin{aligned}
P_{n,k} &= \text{the number of permutations of } n \text{ things taken } k \text{ at a time} \\
&= \underbrace{n(n-1)\cdots(n-k+1)}_{k \text{ factors}} = \frac{n!}{(n-k)!}
\end{aligned}
$$

$$
\begin{aligned}
C_{n,k} &= \left(\begin{array}{c} n \\ k \end{array} \right) = \left(\begin{array}{c} n \\ k, n-k \end{array} \right) \\
&= \text{the number of combinations of } n \text{ things taken } k \text{ at a time} \\
&= \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{P_{n,k}}{k!} = \frac{n!}{k!(n-k)!}.
\end{aligned}
$$

Binomial theorem:

$$(x+y)^n = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) x^k y^{n-k}.$$

Trinomial coefficient:

$$\left(\begin{array}{c} n \\ i, j, k \end{array} \right) = \frac{n!}{i!j!k!} \quad (i + j + k = n,\ i \geq 0, j \geq 0, k \geq 0).$$

For example, $\left(\begin{array}{c} 7 \\ 1, 2, 4 \end{array} \right)$ is the number of words that can be spelled by arranging the letters $ABBCCCC$ in all possible ways.

Date: February 20, 2007.

Conditional probability and independence

Multiplication rule for conditional probabilities.

$$\Pr(A|B)\Pr(B) = \Pr(AB), \quad \Pr(B|A)\Pr(A) = \Pr(AB).$$

Independence.

- Two events A and B are *independent* if $\Pr(AB) = \Pr(A)\Pr(B)$.
- Three events A, B and C are *independent* if $\Pr(AB) = \Pr(A)\Pr(B)$, $\Pr(BC) = \Pr(B)\Pr(C)$, $\Pr(AC) = \Pr(A)\Pr(C)$ and $\Pr(ABC) = \Pr(A)\Pr(B)\Pr(C)$.
- If $A_1, \ldots, A_n$ are independent, and $B_1, \ldots, B_n$ is a list of events such that for $i = 1, \ldots, n$, either $B_i = A_i$ or $B_i = A_i^c$, then $\Pr(B_1 \cdots B_n) = \Pr(B_1) \cdots \Pr(B_n)$.

Bayes' Theorem. Given an event A such that $\Pr(A) > 0$ and a partition $S = B_1 \cup \cdots \cup B_n$ such that $\Pr(B_1) > 0, \ldots, \Pr(B_n) > 0$, we have

$$\Pr(B_i|A) = \frac{\Pr(A|B_i)\Pr(B_i)}{\Pr(A|B_1)\Pr(B_1) + \cdots + \Pr(A|B_n)\Pr(B_n)}$$

for $i = 1, \ldots, n$.

Binomial distribution

If events $A_1, \ldots, A_n$ are independent and each has probability p then

$$\Pr(\text{exactly } k \text{ of the events } A_1, \ldots, A_n \text{ happen})$$

$$= \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for } k = 0, \ldots, n.$$

Univariate distributions

Probability functions (p.f.'s). If X is discrete, then the p.f. f_X satisfies

$$f_X(x) = \Pr(X = x), \qquad \sum_{\text{all possible values } x \text{ for } X} f_X(x) = 1.$$

Probability density functions (p.d.f.'s). If X is continuous, then the p.d.f. f_X satisfies

$$\int_a^b f_X(x)dx = \Pr(a \leq X \leq b) = \Pr(a < X < b), \quad \int_{\text{all possible values } x \text{ for } X} f_X(x)dx = 1.$$

Remember... that we always put $f_X(x) = 0$ when x does not belong to the set where the random variable X is supposed to take values.

Uniform distributions. If X is a continuous random variable with p.d.f.

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{for } x < a \text{ or } x > b \end{cases}$$

we say that X is *uniformly distributed* in the interval $[a, b]$.

Distribution functions (d.f.'s). If X is any random variable (continuous or discrete), the d.f. F_X satisfies

$$F_X(x) = \lim_{t\to x^+} F_X(t) = \Pr(X \le x), \quad \lim_{t\to x^-} F_X(t) = \Pr(X < x),$$

$$\lim_{t\to -\infty} F_X(t) = 0, \quad \lim_{t\to +\infty} F_X(t) = 1, \quad x_1 \le x_2 \Rightarrow F_X(x_1) \le F_X(x_2).$$

Bivariate distributions

Probability functions (p.f.'s). If X and Y are discrete, the joint p.f. $f_{X,Y}$ of X and Y satisfies

$$f_{X,Y}(x,y) = \Pr(X = x \,\&\, Y = y), \qquad \sum_{\substack{\text{all possible values } x \text{ of } X \\ \text{all possible values } y \text{ of } Y}} f_{X,Y}(x,y) = 1.$$

Probability density functions (p.d.f.'s). If X and Y are continuous, the joint p.d.f. $f_{X,Y}$ of X and Y satisfies

$$\Pr((X,Y) \in A) = \int\int_A f_{X,Y}(x,y)dxdy,$$

$$\int\int_{\text{region where } (X,Y) \text{ may take values}} f_{X,Y}(x,y)dxdy = 1.$$

Remember... ... that we always put $f_{X,Y}(x,y) = 0$ if the pair (x,y) does not belong to the set in the (x,y)-plane where the pair (X,Y) of random variables is supposed to take values.

Marginal distributions and independence

Discrete case. If X and Y are discrete, and we are given the joint p.f. $f_{X,Y}$, then the marginal p.f. f_X (which is just the p.f. of X) is given by

$$f_X(x) = \sum_{\text{all possible values } y \text{ of } Y} f_{X,Y}(x,y),$$

and the marginal p.f. f_Y (which is just the p.f. of Y) is given by

$$f_Y(y) = \sum_{\text{all possible values } x \text{ of } X} f_{X,Y}(x,y).$$

Continuous case. If X and Y are continuous, and we are given the joint p.d.f. $f_{X,Y}$, then the marginal p.d.f. f_X (which is just the p.d.f of X) is given by

$$f_X(x) = \int_{\text{all possible values } y \text{ of } Y} f_{X,Y}(y)dy,$$

and the marginal p.d.f. f_Y (which is just the p.d.f of Y) is given by

$$f_Y(y) = \int_{\text{all possible values } x \text{ of } X} f_{X,Y}(x,y)dx.$$

Independence. If X and Y are both discrete and

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

for all x and y without exception, then we say that X and Y are *independent.* If X and Y are continuous, and

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

without exception[1] then we say that X and Y are *independent.*

[1]Well, actually, hardly ever.